Dempster-Shafer Fusion for Personnel Detection

Application of Dempster-Shafer Theory with Ultrasonic micro-Doppler and PIR Sensors

Brian Maguire and Sachi Desai
Acoustics & Networked Sensors Division
US Army RDECOM-ARDEC
Picatinny, NJ, United States of America
brian maguire1@us.army.mil sachi.desai@us.army.mil;

***Abstract*—This paper proposes an approach to the fusion of multi-modal sensor data for the purpose of personnel intrusion detection. The focus is on using low cost non-imaging sensors for applications such as border crossings where issues of rapid deployment *and* power consumption are prevalent. The main challenge of fusing data from such sensors lies in the wide variation of granularity of classification that they may provide. While some sensors may provide detailed characteristics of the motion in a scene and therefore a very fine classification, others may only provide simple alerts and little detail. In order to fuse data from a wide range of sensors that are often designed for disparate applications, an approach based on the Dempster-Shafer theory of evidence is used. The implicit handling of uncertainty and ambiguous propositions leads to a convenient hierarchical approach that can represent data from numerous sensor modalities.**

Keywords-Dempster-Shafer; micro-Doppler; PIR; Personnel Detection;

I. INTRODUCTION

The Dempster-Shafer (D-S) mass function is used in effect as a common representation of the heterogeneous sensor data. In order to cast each data source in this form, first the raw data is reduced to points in a multi-dimensional feature space specific to each sensor. From there, an approach is outlined that uses a distance metric in the feature space to assign mass to each state in the class hierarchy. This hierarchy begins with the full frame of discernment which represents complete uncertainty. From there it proceeds as an n-array tree broken down in to further subclasses until the finest granularity of classification for the specific sensor is reached. For an input point to be classified mass is assigned iteratively down the tree. In doing so, two key steps are taken. First, the uncertainty is estimated as a function of the ratio of the distance between the two closest child nodes. If the input point is deemed equidistant from the child nodes, there is a great deal of uncertainty and the mass function should reflect that. On the other hand, significant disparity indicates a much greater likelihood of one subclass. This distinction leads to the second step, where any mass not assigned to uncertainty is split between the child nodes as a function of the ratio of their distances. The final result is a representation of the likelihood of each singleton class, as well as all unions of these classes representing uncertain states. These D-S mass functions can now be fused using Dempster's rule of combination, and classification rules can be derived to provide a more robust singular solution.

The preceding approach is derived with simulated data, and subsequently demonstrated on two sensor modalities: an ultrasonic micro-Doppler sensor and a PIR profiling sensor. The ultrasonic sensor is able to extract human motion by identifying the periodicity of a human walker's gait in the sensor field of view. The sensor can distinguish between a human, an unknown object in the scene, and background ambience. On the other hand the profiling sensor is capable of distinguishing a horse from a human. The sensor forms a 2-D image of height versus time, and from this the orientation and eccentricity of the object are estimated and matched to known distributions of human and horse profiles. These two sensors illustrate the approach on differing hierarchies of class representations, and an example is given to show the fusion of the two data types. The paper concludes with a discussion of the results and in particular future directions of the work and how enhanced data sets could aid the evaluation of the approach.

II. FUSION BACKGROUND

The method outlined herein is developed based on the Dempster-Shafer theory of evidence. The implicit handling of uncertainty and ambiguous classes in this approach are particularly useful in applications of multi-modal sensor fusion. This approach allows combination of sensor outputs that effectively classify different sets of objects with varying levels of granularity.

Much like in a discrete Bayesian approach, Dempster-Shafer implementation begins by defining the frame of discernment ; the set that contains all possible outcomes ω_q. A mass function assigns belief in each of these outcomes, but unlike a Bayesian probability mass function, mass can also be assigned to unions of outcomes representing uncertainty between them. In this context, the outcomes are specific actors or actions detectable in a scene such as a vehicle moving or a person walking, and any assignment to a union of outcomes indicates uncertainty in performing a finer classification.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
JUL 2012		**00-00-2012 to 00-00-2012**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Dempster-Shafer Fusion for Personnel Detection: Application of Dempster-Shafer Theory with Ultrasonic micro-Doppler and PIR Sensors	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **US Army RDECOM-ARDEC,Acoustics & Networked Sensors Division,Picatinny,NJ,07806**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Presented at the 15th International Conference on Information Fusion held in Sinapore on 9-12 July 2012. Sponsored in part by Office of Naval Research and Office of Naval Research Global.

14. ABSTRACT
This paper proposes an approach to the fusion of multi-modal sensor data for the purpose of personnel intrusion detection. The focus is on using low cost non-imaging sensors for applications such as border crossings where issues of rapid deployment and power consumption are prevalent. The main challenge of fusing data from such sensors lies in the wide variation of granularity of classification that they may provide. While some sensors may provide detailed characteristics of the motion in a scene and therefore a very fine classification others may only provide simple alerts and little detail. In order to fuse data from a wide range of sensors that are often designed for disparate applications, an approach based on the Dempster-Shafer theory of evidence is used. The implicit handling of uncertainty and ambiguous propositions leads to a convenient hierarchical approach that can represent data from numerous sensor modalities.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a REPORT **unclassified**	b ABSTRACT **unclassified**	c THIS PAGE **unclassified**	**Same as Report (SAR)**	**6**	

Two quantities that are useful in generating classification rules based on D-S mass functions are belief and plausibility. Belief is computed as the sum of all outcomes that directly support a given outcome, and can be viewed as an upper bound on the likelihood of said outcome. Plausibility on the other hand is computed as the sum of all outcomes that do not explicitly contradict a given outcome. This serves as a lower bound on the outcome likelihood.

$$Bel(A) = \sum_{B \subseteq A} m(B) \tag{1}$$

$$Pls(A) = \sum_{A \cap B \neq \emptyset} m(B) \tag{2}$$

In this implementation, it is assumed that each sensor produces a mass function in the presence of some input. The mass functions from individual sensors can be combined by Dempster's rule of combination, given in (3). By summing the mass over intersections between the two sensors, mass assigned to more general outcomes is being incorporated to support those that are more specific.

$$m(A) = \frac{\sum_{B \cap C = A} m_1(B) m_2(C)}{1 - \sum_{B \cap C \neq \emptyset} m_1(B) m_2(C)} \tag{3}$$

III. Mass Function Assignment

The approach for automated generation of mass functions from sensor feature data is illustrated by way of an example using simulated data. Figure 1 illustrates the hierarchy of classes in the simulated data set. This hierarchy is designed to simulate a sensor that discerns two levels of classification. For example, classes A and D may represent human and vehicle respectively. Classes B and C in turn would represent specific vehicle types.

Figure 2 shows the generated simulated data. The data is normally distributed in a two dimensional feature space with varied means and unit variance. Consider these points as training samples used to build a model from which mass functions can be derived. Note that no samples are explicitly labeled as class D. All points belonging to sub-classes B and C will be used to define the prior statistics of class D.

In order to proceed, a suitable distance metric must be defined to determine the distance from an input point to each class. For this initial implementation, the Gaussian log likelihood $L(x;\mu_i,\sigma_i^2)$ will be used to generate the vector d containing class distances. Any alternate metric can be chosen to best suit real world data sets.

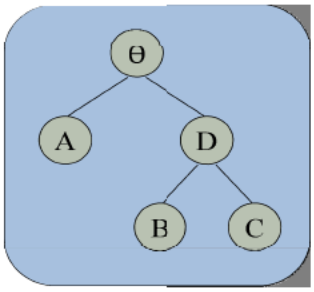

Figure 1. Hierarchy of classes in simulated data

To begin assigning mass to all states in the hierarchy, consider that all mass must sum to 1. Therefore, at the root node □, mass totaling 1 must be split between the left and right sub-trees, with any amount "left behind" representing our uncertainty between the two. Clearly the relative likelihoods of classes in each sub-tree should drive the split, but this fails to account for an uncertainty. To overcome this, the assignment to uncertainty will be considered first as the ratio of the minimum to maximum likelihoods of child nodes as stated in equation (4). In this formulation, when the children are equally likely, the majority of mass will be assigned to total uncertainty. Conversely, when one is far more likely than the other, there is little uncertainty. A scale factor is included to allow some mass to be distributed down the tree even in the case of equally likely sub-trees. The remaining mass is split between the child nodes as in equation (5), where the ratio of distance to the current child to the sum of all child distances scales the assignment.

$$m(\theta) = \alpha \frac{\min(L(x))}{\max(L(x))} \tag{4}$$

$$m(\omega_i) = (1 - m(\theta)) \frac{L(x;\mu_i,\sigma_i^2)}{\sum_j L(x;\mu_j,\sigma_j^2)} \tag{5}$$

This process can now be repeated iteratively down the class hierarchy, with some mass less than 1 remaining in each node to be split amongst the child nodes. Figure 3 presents this approach in algorithmic form. The final output is a mass function that sums to 1 and assigns mass based on the likelihood in and uncertainty amongst each outcome.

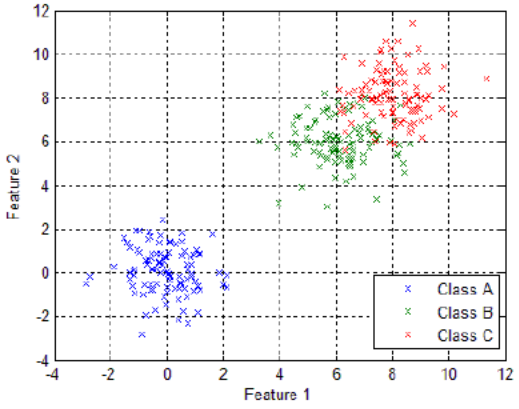

Figure 2. Simulated data

Algorithm Iterative Mass Computation
Input:
- d : vector contains distance from input to each class
- T_i : sub tree tree rooted at node i, initially T
- r_mass: mass at root node, initially 1

compute_ds_mass(d,T,r_mass)
if (!children)
  mass(i) = r_mass
else
  mass(i) = r_mass*exp(-α*max(d)/min(d))
  for j = 1:length(children)
    mass(j) = (1-mass(i))*d(j)/sum(d)
    compute_ds_mass(d,T_j,mass(j))
  end for
end if

Output:
- mass: vector contains mass assigned to all nodes

Figure 3. Algorithm for computing mass function

Consider now a set of input points ranging from -2 to 12 along both feature axes. Figure 4 shows the mass assigned to each class in the hierarchy for all points in this range. All points at the same 2-D location across all five images comprise a computed mass function for a single input point. As is to be expected, the mass peaks for singleton classes are at the mean value of feature space of the training samples of that particular class. For the unions of classes representing uncertainty, the mass peaks at points roughly equidistant between the child classes.

IV. Example Fusion Of Real Sensor Data

In order to illustrate the above approach and the computation of a fused solution, an example is presented. Utilizing data generated by two real-world sensors: an ultrasonic micro-Doppler sensor and a PIR profiling sensor. Both are capable of a basic binary detection when signal energy rises above a certain threshold; beyond that their classification capabilities differ. The profiling sensor can classify between a human and a horse, while the ultrasonic can classify a human but cannot explicitly classify an animal. This difference in granularity is illustrated in Figure 5.

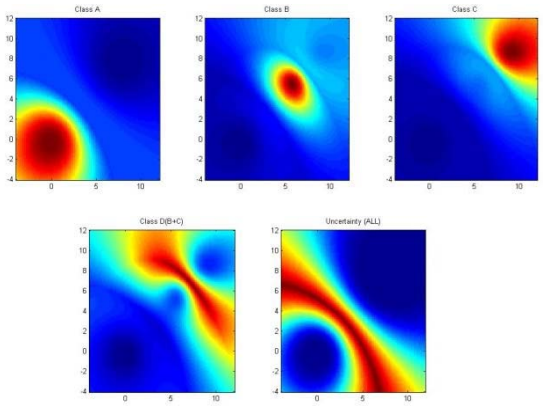

Figure 4: Mass assignments from simulated data

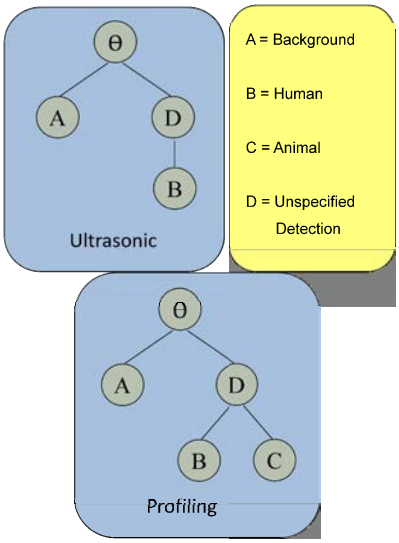

Figure 5. Hierarchy of classes in ultrasonic and profiling data

A. *Ultrasonic Microdoppler Sensor*

The ultrasonic micro-Doppler sensor used here consists of a send and receive channel. The send channel emits a 40 kHz carrier wave, allowing motion in the scene to be characterized by the modulation in frequency sensed at the return channel. In particular, the periodicity of a human gait can be extracted fairly robustly if an actor is walking towards or away from the sensor.

In order to extract meaningful features from the return signal, it is first demodulated. Then a low pass filter is used to remove the carrier signal and noise outside the relevant band. This signal is then Hanning windowed into 4096 sample segments, equivalent to roughly 25 msec at a 160 kHz sampling rate. The FFT of each window is computed and the energy below the LPF cutoff is summed. This step effectively yields a down sampled version of the original signal. The first characterized feature is simply the mean value of this summed energy signal across a 100 sample window, corresponding to approximately 1.2 seconds. This feature provides the basic ability to detect motion of any sort in the scene. The second feature is obtained by computing the autocorrelation of this windowed signal. The main peak of the autocorrelation approximates the main period of motion of the scene, and is effective at classifying a human walking versus random motion. Figure 6a-c on the following page shows the steps of this process: the spectrogram of the down sampled and filtered signal, the summed energy signal, which roughly traces the periodicity in the spectrogram, and finally the autocorrelation and the computed peak.

B. *PIR Profiling Sensor*

The profiling sensor used in this paper consists of a linear array of 11 nodes. The sensor in effect produces a two dimensional image with the axes being time and height. This image is then post processed to produce profiles of objects,

which can be useful in discerning objects of different dimensions and orientations.

Given a sample of the profiling data windowed in time, it is first necessary to subtract the values in each row from the mean of that row. This procedure produces a normalized image, and the standard deviation of the pixel intensity values serves as the first feature to discern an object in the scene.

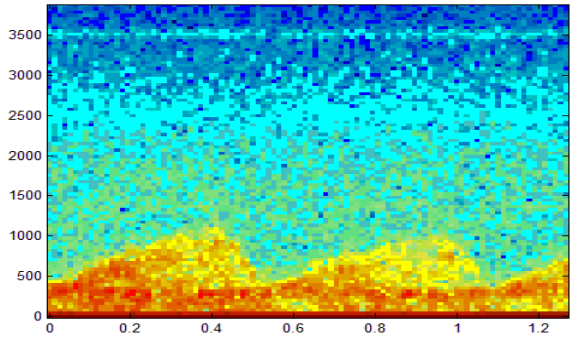

Figure 6a. Spectrogram of the down sampled and filtered signal

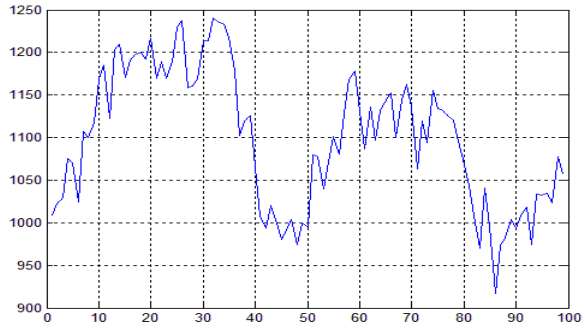

Figure 6b. Summed energy signal

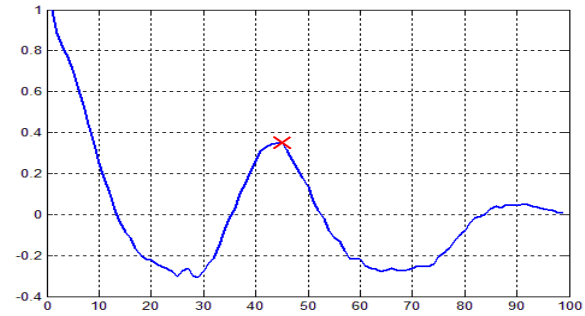

Figure 6c. Autocorrelation and the computed peak

Next the image is converted to a binary image by applying a threshold to the pixel intensities. The coordinates of all pixels that exceed the threshold are arranged in a two column matrix and the covariance matrix is computed. The eigen decomposition of this covariance matrix is performed. As described in [9], the objects orientation and eccentricity are estimated from these values. The orientation is estimated as the angle from the image x-axis to the eigenvector associated with the largest eigenvalue. The eccentricity is estimated by the ratio of the two eigenvalues.

Figure 7a-c illustrates the steps of the feature extraction for a human sample. This figure includes the original image, the normalized image, and the threshold image. The computed major and minor axes of the image are overlaid on the threshold image for clarity. Note the image axes are not of equal scale, so although orthogonal axes are computed, they do not appear as such.

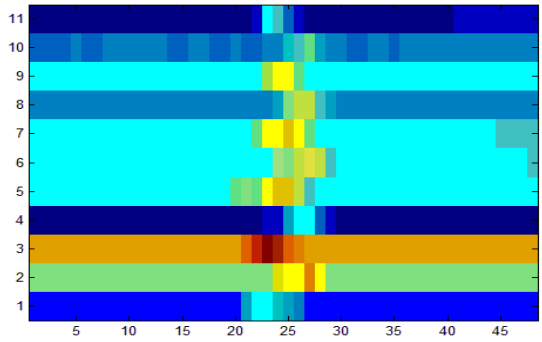

Figure 7a. Original image from profiling sensor

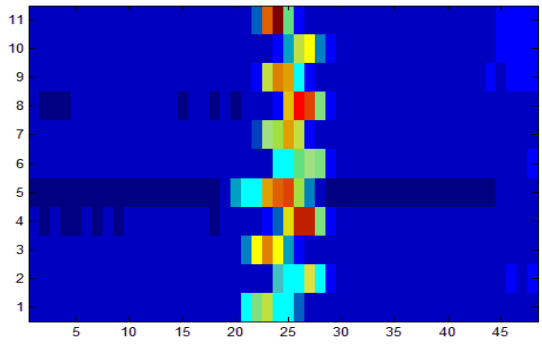

Figure 7b. Normalized image of profiling data

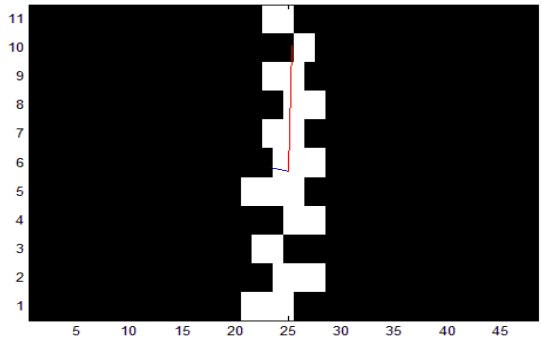

Figure 7c. Threshold image from profiling data for feature extraction

C. Realization

The realization of the D-S based approach is greatly simplified by the use of a class membership matrix, *C*, for all computations. Similar to a graph adjacency matrix, the class membership matrix takes on binary values to indicate to which parent classes each child class belongs, and conversely which classes are child classes of each parent. The matrix is populated as in equation (6).

$$C(i,j) = \begin{cases} 1, \omega_i \subseteq \omega_j \\ 0, \omega_i \not\subset \omega_j \end{cases} \quad (6)$$

The parent classes to which a child belongs can now be determined by simply searching for 1 in the corresponding column. Similarly, the children of a class are found by searching the corresponding row. For example, each iteration of the algorithm depicted in Figure 3, the matrix is searched first for any columns with a sum of 1, indicating a top level node. The corresponding row is then searched to determine its children. The matrix is pared down and passed through each iteration until there are no child nodes left. This process is also used to ease the computation of belief and plausibility, as well as fused masses. For example, belief can be computed by simply multiplying the matrix by a vector of computed masses since each row selects a proposition that directly supports the corresponding class. Plausibility can be computed by performing a logical OR of the same row and column, selecting all indices in the mass function that do not contradict the corresponding class.

D. Sample Fusion

Using the aforementioned approaches, raw data from each of the two sensors is converted in to a feature space representation. The ultrasonic data set contains 88 samples of human motion, 43 samples of non-human motion, and 81 samples of background with no motion. The profiling data set contains 92 samples of human motion, 43 samples of horse motion, and 102 samples of no motion. Figure 8a-b shows the distribution of training samples for both sensors in their respective feature space.

It should be noted that although the data here is of similar objects, it was not collected concurrently and therefore does not provide a detailed assessment of improvement in classification accuracy. The goal here is only to illustrate the ability of the proposed approach to combine information from two disparate sensors, which provide different levels of classification.

To provide an example of the fusion of these two sensors, one human sample is left out of each sensor's data set. The prior distributions for all classes, including uncertainty, are computed using the remaining samples. The left out samples are treated as a coincident test sample, and from these each sensor model produces a D-S mass function. These mass functions are fused using Dempster's rule of combination. Given in Table 1 is the resulting individual mass function computation of each sensor and the fused result of the two sensors. Table 1 follows the class hierarchy established in figure 5.

In this example, the profiling sensor has substantial uncertainty in the class of the object. While a simple classification rule based on largest individual mass assignment would still assign the class label B for human, there is also significant mass assigned to both uncertainty and background. Clearly the profile of the human is narrow in time, which leads to small peaks. The small peaks cause some confusion with the background when considering the standard deviation of the pixels. The ultrasonic sensor similarly assigns the majority of mass to class B human, but in this case assigns substantial mass to class D. While it is near certain that motion has been detected in the scene, it is unclear whether the motion can be specifically ascribed to a human. Fusing these two results produces a much more convincing mass function. In this case, nearly all mass is assigned to support class B human.

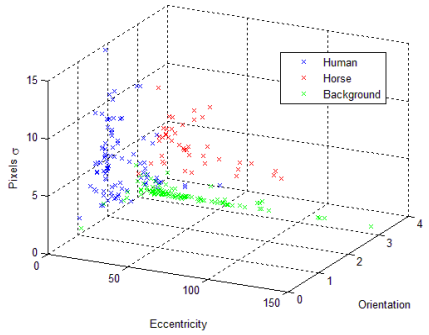

Figure 8a. Feature space plots of profiling training data

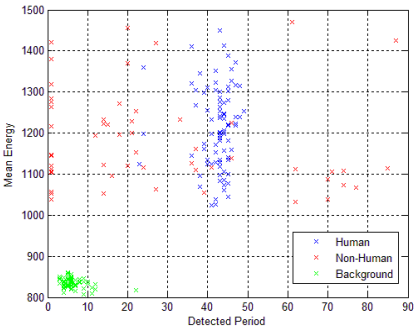

Figure 8b. Feature space plot of ultrasonic micro-Doppler training data

Table 1. Results of mass computation and fusion

Sensor	Class				
	A	B	C	D	
Profiling	0.1899	0.5667	0.0415	0.0007	0.2012
Ultrasonic	0.0150	0.5290	---	0.4560	0.0001
Fused	0.0075	0.8505	0.0242	0.1177	0.0001

V. SUMMARY

Herein we provide an application of the Dempster-Shafer theory method for personnel detection. The Dempster-Shafer theory provides the capability of fusing orthogonal data from an ultrasonic micro-Doppler and PIR sensors. Utilizing two sets of real-world data from aforementioned sensors that were collected separately we complete an investigation of the benefits of sensor data fusion. The application of the outlined approach provides a hierarchal approach to classification / discrimination through fusion of the disparate information resulting in a series of solutions with a greater confidence in comparison to a standalone sensor solution. In our analysis several areas of improvement result specifically the capability to have multiple classes and multiple confidence states for given problem. The utilization of multiple classes afforded by the Dempster-Shafer theory increases the robustness and quality of the information from the given suite of sensors.

ACKNOWLEDGMENT

The authors would like to acknowledge Dr. James Sabatier and Dr. Alexander Ekimov of National Center of Physical Acoustics at the University of Mississippi.

REFERENCES

[1] Valcourt, David. United States . TRADOC Pamphlet 525-66. Washington D.C.: U.S. Army, 2008.

[2] Koks, Don, and Subhash Challa. Australia. TRADOC Pamphlet 525-66. Edinburgh: DSTO, 2005.

[3] Denoeux, Thierry. "A Neural Network Classifier Based on Dempster-Shafer Theory." IEEE Transactions on Systems, Man, and Cybernetics. 30.2 (2000): 131-150.

[4] Zouhal, Lalla, and Thierry Denoeux. "An Evidence-Theoretic k-NN Rule with Parameter Optimization." IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS. 28.2 (1998): 263-271

[5] Llinas, J, and D Hall. "An Introduction to Multi-Sensor Data Fusion." IEEE International Symposium on Circuits and Systems. Monterey, IEEE. 1999. Print.

[6] D. L. Hall and J. LLinas. Multisensor Data Fusion. In D. L. Hall and J. LLinas, editors, Handbook of Multisensor Data Fusion, chap. 1, pp. 1–1 – 1–10. CRC Press LLC, 2001.

[7] Shilov, Georgi. Linear Algebra. Revised English Edition., New York: Dover Publications, Inc., 1977.

[8] Li, Lingling, Zhigang Li, Meng Wu, and Chuntao Zhao. "Decision-making Based on Dempster-Shafer Evidence Theory and its Application in the Product Design." Applied Mechanics and Materials. 44.47 (2010): 2724-2727.

[9] Gonzalez, Rafael, Richard Woods, Digital Image Processing. Upper Saddle River, NJ: Prentice Hall, 2002

\